17-14 Artificial Intelligence Technology

Status: New **Issued:** R.O. 20XX.XX.XX-XXXX

Rationale

The Toronto Police Service (Service) is continually seeking to provide effective policing through the prudent adoption of new technologies while ensuring transparency, service provision in accordance with legislative requirements and the interests of the community. The Service strives to maximise its ability to protect the public while ensuring the legitimate use of all data collected, the preservation of privacy rights, and engaging in continuous evaluation of privacy and ethical implications of new technologies.

Artificial intelligence (AI) technology can support the Service's mission to deliver police services, in partnership with our communities, to keep Toronto the best and safest place to be. Technology can provide opportunities to improve efficiency and customer service, resolve complex issues, ensure business continuity, and provide assistance in performing repetitive tasks; however, AI technology also carries risks that must be identified and mitigated.

The purpose of this Procedure, in compliance with Toronto Police Services Board (Board) policy, is to provide a framework for the acquisition and use of AI technology, using a measured approach to risk assessment, costs, and benefits related to the use of this technology.

Procedure

For the purpose of this Procedure, Artificial Intelligence refers to any and all technology within the AI spectrum of applications.

Thoughtful consideration of the benefits and risks of obtaining and deploying new AI technology is necessary to ensure the technology supports public trust in the Service, community safety and sense of security, individual dignity, and the equitable and effective delivery of policing services. The Service seeks to remain current in its ability to effectively investigate alleged criminal behaviour while ensuring the protection of individual liberty and operating in accordance with the law. All new AI technologies shall be assessed to avoid amplification of marginalization or discrimination.

Ethics and Transparency

The Service is committed to continuously ensuring bias is avoided in all new AI technology being considered. The ethical, legal and community impacts of all new AI technology will be thoroughly reviewed to ensure legal compliance and transparency to the public.

VISIT…

Risk Categorization

For the purposes of this procedure, a detailed screening process will be used to apply a risk category of "Sensitive" or "Not Sensitive" to any new AI technology proposed for Service use. Prior to any AI technology being approved for Service use, it will be assigned a risk level consistent with the categories outlined in the Board Policy *Use of Artificial Intelligence Technology* (Board Policy).

"Sensitive" refers to AI technology that has the potential to impact rights, privacy, marginalization or discrimination. This includes any AI technology within the Board Policy risk categories *low, moderate, high,* and *extreme*. As such, all AI technology that falls within the Sensitive category requires a review to determine its purpose and attributes in order to evaluate the impacts on individual rights and public security/safety.

There are other AI technologies specifically tasked to assist manual processes, with no implications to rights, privacy, marginalization, discrimination, or to the public. The purpose of this technology is to relieve human resources to work more efficiently by assisting with a manual and/or tedious processes. For example, this could include software to assist with inventory management of uniform clothing, or a computer network program to detect and defend against malware attacks. These AI technologies may be screened as "Not Sensitive" and includes any AI technology within the *minimal* risk category, as outlined in the Board Policy.

All AI technology will require appropriate review for approval before Service use.

Review and Approval Process

Any AI technology proposed for Service use shall be reviewed through the formal approval process to ensure the Service's mandate is upheld and ethical implications of the technology are thoroughly considered. This review is conducted by the Artificial Intelligence Technology Committee (AITC).

The AITC membership consists of representatives from the following areas:

- Committee Chair – Chief Information Officer - Information & Technology Command (ITC), or designate
- Office of the Chief
- Legal Services (LSV)
- Strategy Management (STM)
- Information Technology Services (ITS)
- Information Management (IM)
- Information Privacy & Security
- Business Relationship Management (BRM)
- Ad Hoc members, as necessary, dictated by the individual project scope or type of technology under consideration

The review will address the intended benefit that an AI technology will deliver and will consist of the following:

- Business Initiative Review (BIR)
  - to validate the business need for the technology including how it applies to the Services' mandate and provision of services
- Artificial Intelligence Pre-Assessment Screening (TPS 209)
  - to identify the details of the AI technology to determine if it meets the initial Risk Categorization of "Sensitive"
- AITC Initial Assessment

  - to identify and evaluate risks, and to determine alignment with the Service's mandate and business plan
  - to identify details about the situations in which it is proposed to be used
  - to decide if the AI technology should proceed to the next step of the process
  - to decide the scope of stakeholder consultation
- Stakeholder Consultation
  - as necessary, consultations with internal, external agency, and community stakeholders for feedback and awareness
- Testing and Verification
  - to analyze and verify the details of the technology and infrastructure
- AITC Post-test Review
  - to review the test and verification results, and share with stakeholders as necessary
  - to determine if more testing is required
  - to determine if the AI technology should proceed to the next step of the process
- Artificial Intelligence Assessment (AIA) and Privacy Impact Assessment (PIA)
  - to review the mechanical, technical, social, and privacy components of the technology.
  - All AI technology applications require both an AIA and a PIA.
- Final Review
  - conducted by the AITC, in consultation with the Chief of Police (Chief), Command, the Board and external agencies, as necessary.

Member

1. While conducting police business shall only use AI technology authorized by the Service.

   ➢ *Note: If the AI technology is not Service authorized, members shall not use it in the course of their duties.*

2. When using data related to Service authorized AI technology, members shall only use it
   - for which it was originally intended and authorized for Service use, and
   - in compliance with the Standards of Conduct and incident specific Service Procedures

3. When identifying a business or operational need that can be or may be assisted by new AI technology, shall engage BRM via email to
   - identify any existing and effective Service approved technology that can perform the desired function(s) and, where no such technology exists,
   - prepare a Business Initiative Request (BIR) for the purpose of initiating the AI technology review and approval process.

4. When designated the project lead/Service sponsor for a proposal to obtain new AI technology shall
   - complete a TPS 209 and submit to BRM
   - perform tasks as directed by the AITC, including
     – documenting the project details and scope
     – reporting on project progress and technology testing results
   - create a project team, as necessary

Member – Business Relationship Management (BRM)

5. When in receipt of a proposal for new AI technology to assist a business or operational need shall maintain a central repository for tracking all requests, including:
   - the proposed technology name and vendor
   - the project lead/Service sponsor

- the status of the request through the review process
- reasons for approving/denying the request throughout the review process

6. When in receipt of a proposal for new AI technology to assist a business or operational need for which there is existing and effective Service approved technology to perform the desired function(s) shall
   - advise the requesting member of the existing technology
   - deny the request for new AI technology

     ➢ *Note: Denial of a request does not preclude the AI technology from being revisited at a future date. Funding, technological advances, and/or organizational/community readiness may give reason to re-evaluate the AI technology.*

7. When in receipt of a proposal for new AI technology to assist a business or operational need for which there is not an existing and/or effective Service approved technology to perform the desired function(s) shall
   - consider any alternate options to achieve the same goal as proposed by the new technology
   - confirm the primary stakeholder to act as project lead / Service sponsor
   - ensure the project lead/Service sponsor completes the TPS 209
   - prepare a BIR
   - forward the completed BIR and TPS 209 to the Manager – BRM

Manager - Business Relationship Management (BRM)

8. Upon receipt of the BIR and TPS 209 regarding a proposal for new AI technology shall ensure
   - the TPS 209 form has been completed in full
   - the details of the proposed AI technology have been properly documented and catalogued for future reference
   - both forms are forwarded to the AITC
   - a member is assigned to maintain the central repository for AI technology requests

Artificial Intelligence Technology Committee (AITC)

9. Shall establish
   - the AITC mandate
   - the AITC policy outlining information to ensure a thorough review is conducted for all new AI technology

     ➢ *Note: this could include: terms of service; independent third party audits; purpose and function of the AI technology; circumstances/situations in which the AI technology will be used; context for which the AI technology will be used (i.e. in a silo or alongside other technologies/practices, cumulative benefit/risk when used in silo/alongside other technologies/practices, etc.).*

10. Upon receipt of a BIR and TPS 209 for new AI technology shall
    - determine the scope of intended use
    - determine the technology's likelihood of impact on the public
    - evaluate the TPS 209 and classify the technology as
      – 'Sensitive' when 'Yes' is indicated in response to any of the risk escalator questions
      – 'Not Sensitive' when 'No' is indicated in response to all of the risk escalator questions
    - evaluate the risk mitigators in the TPS 209
    - determine if the opinions of subject matter experts (both internal and/or external) would be beneficial, including Legal Services and the Crown Attorney

- determine if community consultation and/or representation is needed
- consider any alternate options to achieve the same goal as proposed by the new technology
- comply with the Board Policy, as required
- determine whether or not to continue with the review process for the new AI technology and, when discontinuing the review (i.e. denying the request for new AI technology),
  – advise the project sponsor/Service sponsor of the reasons for denying the request
  – advise BRM, for the purpose of maintaining the central repository of AI requests

  ➢ *Note: Denial of a request does not preclude the AI technology from being revisited at a future date. Funding, technological advances, and/or organizational/community readiness may give reason to re-evaluate the AI technology.*

11. When continuing with the review process for new AI technology shall
    - engage the project lead/Service sponsor to document the scope, details and progress of the project
    - obtain approval from the CIO to engage in AI technology testing, when testing is necessary

      ➢ *Note: Personally identifying information and operational data shall not be used in testing environments.*

    - advise BRM for the purpose of maintaining the central repository for AI requests

12. Upon receipt of the AI technology testing results, shall conduct a post-test review that includes
    - re-evaluating the categorization of the technology
    - evaluating if the technology fulfills the original business need
    - evaluating if the technology functions as intended
    - determining if further testing is required
    - determining if the technology will continue to proceed through the review process

    and

    when determining to proceed with the AI technology review process, the AITC Chair shall forward the project details and test findings to
    - Analytics & Innovation to conduct an Artificial Intelligence Assessment (AIA), and
    - Information Privacy & Security to conduct a Privacy Impact Assessment (PIA)

13. Upon review of the completed AIA and PIA, Board Policy and all information obtained in the review process, shall
    - determine whether or not to continue with the review process for the new AI technology
    - when discontinuing the review (i.e. denying the request for new AI technology),
      – advise the project sponsor/Service sponsor of the reasons for denying the request
      – advise BRM, for the purpose of maintaining the central repository of AI requests

      ➢ *Note: Denial of a request does not preclude the AI technology from being revisited at a future date. Funding, technological advances, and/or organizational/community readiness may give reason to re-evaluate the AI technology.*

    - when continuing with the review, shall
      – assign an initial risk category to the technology, as described in the Board Policy
      – prepare the conditions for use
      – submit all review documents to the CIO
      – make a request to the CIO to engage the Chief and Command and, as necessary, the Board

14. When new AI technology is approved for Service use by the Chief, the AITC shall

- advise BRM, for the purpose of maintaining the central repository of AI requests
- when AI technology needs to be purchased, advise the project lead/Service sponsor to engage BRM
- ensure an evaluation and re-assessment of the technology takes place post-deployment, per the Board Policy

Manager – Analytics & Innovation

- Upon receipt of project details and test findings from the AITC regarding new AI technology shall complete an AIA and submit to the AITC. At minimum, the AIA shall include
  - a functional description of the AI and the processes/decisions it is acting upon
  - description of any bias and limitations, and methods to mitigate
  - results of any testing or reviews
  - a description of the data used by the application, or in lieu of a description may provide a link to the website when the data is publically available
  - an evaluation of the relevant data security, with mitigation and assessment recommendations

Manager - Information Privacy & Security

15. Upon receipt of project details and test findings from the AITC regarding new AI technology shall complete a PIA and submit to the AITC.

Chief Information Officer – Information & Technology Command

16. Upon receipt of a recommendation from the AITC to continue the review process for new AI technology shall ensure consultation with the Chief and Command, with escalation to the Board according to the Board Policy.

17. Maintain a publicly available listing of approved technologies that includes the following details
    - product name
    - product vendor
    - boundaries of use
    - identified risks
    - risk mitigation strategies
    - a record of reviews undertaken by the Service or by third party organizations

      ➢ *Note: Where the product is covert in nature the product and vendor may be omitted. Where public disclosure of the product or details of the product may impact economic, intellectual property, or other interests, such information may be omitted pursuant to sections 10 and 11 of the Municipal Freedom of Information and Protection of Privacy Act"*

Appendices

Appendix A – Artificial Intelligence Technology Acquisition Approval Process

Supplementary Information

Governing Authorities

Federal: TBD

Provincial: Police Services Act, O. Reg 58/16, Collection of Identifying Information in Certain Circumstances – Prohibition and Duties; Police Services Act, O. Reg 3/99, Adequacy and Effectiveness of Police Services; Police Services Act, O. Reg 265/98, Disclosure of Personal Information; Municipal Freedom of Information and Protection of Privacy Act, R.S.O. 1990, c. M. 56

Municipal: TBD

Other: TBD

Associated Governance

Toronto Police Services Board Policies: Use of Artificial Intelligence Technology; By-law No. 163 Purchasing By-Law.

Service Procedures: TBD

Other: Ethical Decision Making – Standards of Conduct; 1.12 Standards of Conduct – Confidential Information; 1.13 Standards of Conduct – Release of Service Documents; 1.14 Standards of Conduct – Removal of Service File, Record, Exhibit and Property; 1.15 Standards of Conduct – Use of Service facilities & Equipment; 1.19 Standards of Conduct – Use of Computers and Telecommunications.

Forms: TPS 209 Artificial Intelligence Pre-Assessment Screening

Definitions

For the purposes of this Procedure, the following definitions will apply:

Artificial Intelligence (AI) is a science and engineering approach to integrate computer and human behaviour in support of business processes. This refers to the general concept of a non-human program or model that can solve problems and perform sophisticated calculations.

Bias for the purpose of this procedure means

(a) stereotyping, prejudice or favoritism towards some things, people, or groups over others. These biases can affect collection and interpretation of data, the design of a system, and how users interact with a system. Forms of this type of bias include: automation bias, confirmation bias, experimenter's bias, group attribution bias, implicit bias, in-group bias, out-group homogeneity bias.

(b) systematic error introduced by a sampling or reporting procedure. Forms of this type of bias include: coverage bias, non-response bias, participation bias, reporting bias, sampling bias, selection bias. This is not to be confused with the bias term in artificial intelligence learning models or prediction bias

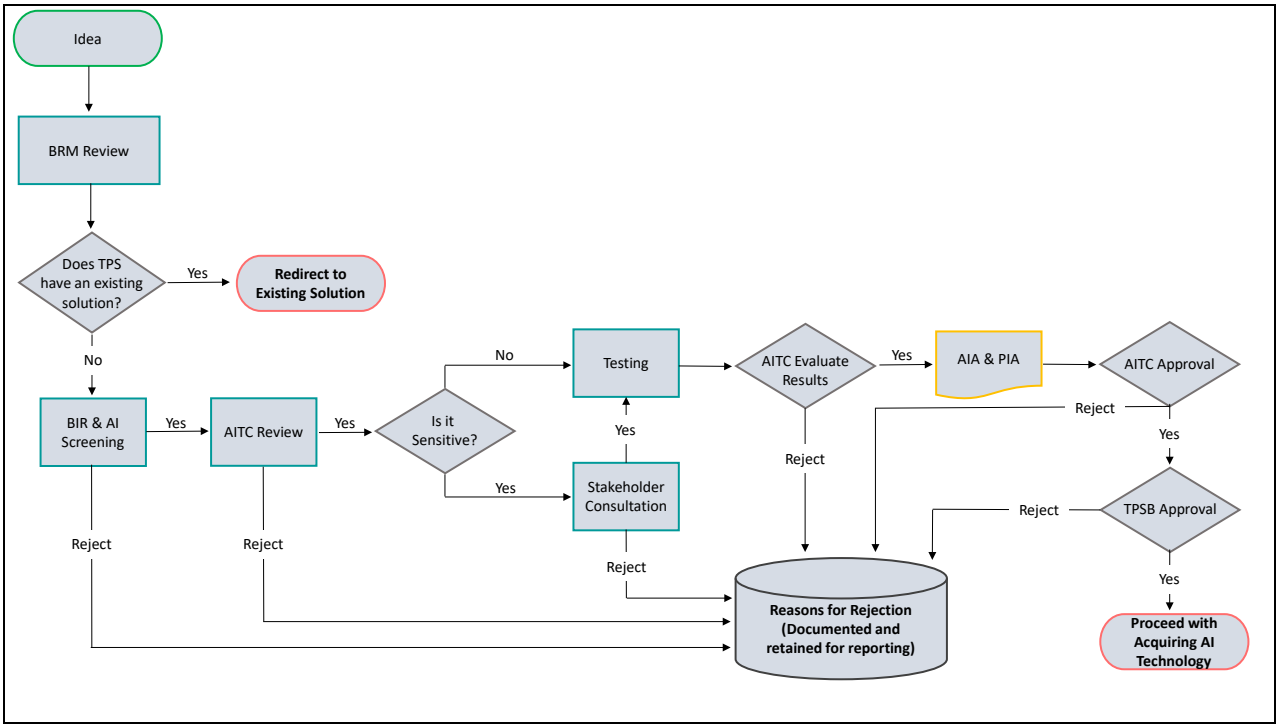
Sensitive (Artificial Intelligence Technology) means Artificial Intelligence (AI) technology that has the potential to impact privacy, marginalization or discrimination. As such, all AI technology that falls within the Sensitive category requires a review to determine its purpose and attributes to evaluate the impacts on individual rights and public security/safety. This includes any AI technology within the following risk categories of the Toronto Police Services Board Policy entitled *Use of Artificial Intelligence Technology*: Low, Moderate, High, and Extreme.

Not Sensitive (Artificial Intelligence Technology) means Artificial Intelligence (AI) technology specifically tasked to assist a manual process, no implications for privacy, marginalization, discrimination, or to the public. The purpose of the technology is to assist a manual and/or tedious processes to relieve human resources to work more efficiently. For example, this could include software to assist with inventory management of uniform clothing, or a computer network program to detect and defend against malware attacks. This includes any AI technology within the Minimal Risk category of the Toronto Police Services Board Policy entitled *Use of Artificial Intelligence Technology*.

17-14 Appendix A

Artificial Intelligence Technology Acquisition Approval Process

Status: New **Issued:** R.O. 20xx.xx.xx–xxxx



Legend:

AI	Artificial Intelligence
AIA	Artificial Intelligence Assessment
AITC	Artificial Intelligence Technology Committee
BIR	Business Initiative Request
BRM	Business Relationship Management
PIA	Privacy Impact Assessment
TPS	Toronto Police Service
TPSB	Toronto Police Services Board

Artificial Intelligence Pre-Assessment Screening

Note: *To be completed prior to procurement, purchase or use of software solutions which contain Artificial Intelligence (AI). Please refer to Procedure 17-14.*

Pre-Assessment Questions to Determine the Initial Risk Categorization

State the purpose of the technology solution and function of the AI application.
(What problem are you trying to solve?)

State the benefit of the technology solution and AI component(s) within.
(How does this help the people we serve?)

Risk Escalators
(Where YES is indicated the solution will be classified "Sensitive" per Procedure 17-14.)

Question	Yes	No	Explanation of Risk
Does the solution contravene the Criminal Code, Charter of Rights, Privacy legislation, or other legislation?	○	○	Solutions must conform to all legislation.
Will the solution affect the level of police resources in a community?	○	○	Solutions which allocate policing resources to communities may reinforce biases and result in unwarranted and increased attention in neighbourhoods.
Will the output of the solution create an immediate policing effect or action?	○	○	If a qualified person is not required to review the results, an improper and consequential action could take place. Results should be considered by qualified personnel before being acted upon.
Does the solution result in indiscriminate or covert monitoring?	○	○	By Collecting and/or analyzing large data sets on individuals and groups, an AI tool may be a form of electronic surveillance which may violate privacy and Charter rights.
Is training data known to have quality issues, control of bias issues that cannot be mitigated?	○	○	AI is very sensitive to bad data. If the data used to train the algorithm is fault or biased, results will also be flawed. The solution must also have protections in place to prevent malicious actions to intentionally achieve negative results.
Will the solution be used to manage evidence?	○	○	Evidence must be handled and protected correctly. Without due diligence, use of AI in this area may cause errors.
Will the solution be used to point to particular persons as suspects in an investigation?	○	○	Use of an AI solution to direct investigators towards person(s) in an indiscriminate fashion may violate rights and legislative requirements.
Does the solution make predictions about behaviour or actions relating to the public?	○	○	Predictions of behaviour can be used to limit an individual's freedoms based on hypothetical events, and must be carefully considered.

Risk Mitigators

(These items mitigate the risk but cannot offset an escalator above without further evaluation.)

Question	Yes	No	Explanation of Risk
Does the solution provide results with additional information to assist the operator?	○	○	If the solution is designed to demonstrate results in a fair and understandable format, it will improve proper evaluation before action is taken.
Will the user be able to identify erroneous or biased results?	○	○	If AI results are shown in a way that the user can identify bias or incorrect results, then risk of failure is mitigated.
Is the solution only to be used under judicial oversight?	○	○	If so, the use of the solution will be documented and required judicial approval before a justice or judge prior to use.
Can systemic biases be identified, measured and countered before impacting the business process?	○	○	Inaccurate AI results can be countered by procedural steps to ensure the bias does not have any impact or effect.
Can the AI/ML capabilities of a solution be restricted or turned off?	○	○	Solutions may have AI which is not central to the functioning of the solution but augments a capability, and can be turned off or limited to select users.
Has a neutral third-party certified the AI/ML solution?	○	○	Previous inspection of the solutions components (data sources, algorithms used, accuracy) can assist with assessment for fair use.
Does the solution have the capability to be audited?	○	○	Auditing allows for a transparent way to see how the solution is used, its effectiveness, and possible indicators of improper use.
Can the solution be monitored for effectiveness, and at what frequency?	○	○	Monitoring can identify errors and accuracy issues to re-evaluate continued use of the solution.
Is this technology widely used for the proposed purpose?* If YES, are there commonly known and accepted best practices for its use? **Include reference documents with the TPS 209 submission when 'YES' is selected.*	○ ○	○ ○	Commonly-known best practices can provide clear guidance on appropriate and acceptable uses, particularly if they have been subject to independent scrutiny.

SUBMIT to Business Relationship Management

DISTRIBUTION: *Original - email to Business Relationship Management*